nature research

Corresponding author(s): Karthik A. Jagadeesh, Kushal K. Dey, Alkes L. Price, Aviv Regev

Last updated by author(s): May 19, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	We used the Harmony python package for single cell batch correction, the scanpy python package for single cell data analysis, the nnet R package, the LDSC package available on github at https://github.com/bulik/ldsc for computing genetic heritability, and the MAGMA gene/gene set prioritization software. Additionally, all custom code developed in this study for analysis of single cell data is available at github at: https://github.com/kkdey/GSSG and https://github.com/karthikj89/scgenetics. This work uses the S-LDSC software (https://github.com/bulik/ldsc) to process GWAS summary statistics as well as S-LDSC software and MAGMA v1.08 (https://ctg.cncr.nl/software/magma) for post-hoc analysis. Code for constructing cell type, disease-dependent and cellular process gene programs from scRNA-seq data and performing the healthy and disease shared NMF can be found at https://github.com/karthikj89/scgenetics (DOI 10.5281/zenodo.6516048). Code for processing gene programs and combining with enhancer-gene links can be found at https://github.com/kkdey/GSSG (DOI 10.5281/zenodo.6513166).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All postprocessed scRNA-seq data (except for Alzheimer's disease; see below) are available through the original publications with PMIDs: 28091601, 33208946, 31316211, 31097668, 31042697, 31348891, 32832598, 31209336, 31604275, 33654293, 32403949, 30355494. Additionally, gene programs, enhancer-gene linking annotations, supplementary data files and high-resolution figures are publicly available online at https://data.broadinstitute.org/alkesgroup/LDSCORE/Jagadeesh_Dey_sclinker. The Alzheimer's disease scRNA-seq data8 is available exclusively at https://www.radc.rush.edu/docs/omics.htm per its data usage terms. This work used summary statistics from the UK Biobank study (http://www.ukbiobank.ac.uk/). The summary statistics for UK Biobank used in this paper are available at https://data.broadinstitute.org/alkesgroup/UKBB/. The 1000 Genomes Project Phase 3 data are available at ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/2013050. The baseline-LD annotations are available at https://data.broadinstitute.org/alkesgroup/LDSCORE/. We provide a web interface to visualize the enrichment results for different programs used in our analysis at: https://share.streamlit.io/karthikj89/scgenetics/www/scgwas.py.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We broadly analyzed across 60 available GWAS data that were picked based on relevance to the scRNA-seq data analyzed.
Data exclusions	We analyzed only autosomes based on pre-established exclusion criteria as seen in Finucane et al 2018.
Replication	Where possible, we replicated our computational results using independent scRNA-seq data sets from the same tissue.
Randomization	We did not allocate samples in experimental groups.
Blinding	There was no group allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging